EXHIBIT P

Environmental and Molecular Mutagenesis 61:66–83 (2020)

Commentary

Quantitative Interpretation of Genetic Toxicity Dose-Response Data for Risk Assessment and Regulatory Decision-Making: Current Status and Emerging Priorities

Paul A. White,[1]* Alexandra S. Long,[2] and George E. Johnson [3]

[1]*Environmental Health Science and Research Bureau, Health Canada, Ottawa, Ontario, Canada*
[2]*Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario, Canada*
[3]*Swansea University Medical School, Swansea, Wales, United Kingdom*

The screen-and-bin approach for interpretation of genotoxicity data is predicated on three false assumptions: that genotoxicants are rare, that genotoxicity dose–response functions do not contain a low-dose region mechanistically characterized by zero-order kinetics, and that genotoxicity is not a *bona fide* toxicological endpoint. Consequently, there is a need to develop and implement quantitative methods to interpret genotoxicity dose–response data for risk assessment and regulatory decision-making. Standardized methods to analyze dose–response data, and determine point-of-departure (PoD) metrics, have been established; the most robust PoD is the benchmark dose (BMD). However, there are no standards for regulatory interpretation of mutagenicity BMDs. Although 5–10% is often used as a critical effect size (CES) for BMD determination, values for genotoxicity endpoints have not been established. The use of BMDs to determine health-based guidance values (HBGVs) requires assessment factors (AFs) to account for interspecies differences and variability in human sensitivity. Default AFs used for other endpoints may not be appropriate for interpretation of *in vivo* mutagenicity BMDs. Analyses of published dose–response data showing the effects of compensatory pathway deficiency indicate that AFs for sensitivity differences should be in the range of 2–20. Additional analyses indicate that the AF to compensate for short treatment durations should be in the range of 5–15. Future work should use available data to empirically determine endpoint-specific CES values; similarly, to determine AF values for BMD adjustment. Future work should also evaluate the ability to use *in vitro* dose–response data for risk assessment, and the utility of probabilistic methods for determination of mutagenicity HBGVs. Environ. Mol. Mutagen. 61:66–83, 2020.

Key words: risk assessment; dose–response analysis; benchmark dose

THE CHANGING PARADIGM FOR INTERPRETATION OF GENOTOXICITY DATA

Genetic toxicity test results have traditionally been employed only for the identification of genotoxic substances, that is, hazard identification. Accordingly, test results are rarely used to assess genotoxic potency; moreover, they are rarely used to determine potency metrics for risk assessment and regulatory decision-making (White and Johnson 2016). This pervasive *screen-and-bin* paradigm is predicated on three false assumptions: (1) genotoxic substances are rare and exposures entirely avoidable, (2) genotoxicity dose–response (DR) functions do not contain a low-dose region that is mechanistically characterized by zero-order kinetics (i.e., a low-dose region within which the response cannot be distinguished from the ever-present background), and (3) genotoxic effects cannot be regarded as *bona fide* toxicological endpoints.

With respect to the first assumption, numerous researchers have documented and discussed the pervasive nature of unavoidable genotoxicants. For example, van der Hoeven et al. (1983) described the mutagenicity of extracts from common foods such as string beans, rhubarb, and lettuce. Additionally, plant-derived natural health products have been shown to elicit mutagenic effects, for example, Black Cohosh and Chinese herbal remedies containing plant material from the genus Aristolochia (Bhalli et al. 2013; Smith-Roe et al. 2018). Smith-Roe et al. (2018) showed that black

Reproduced with the permission of the Minister of Health Canada

Grant sponsor: Health Canada, under the Government of Canada's Chemicals Management Plan.

*Correspondence to: Paul A. White, Environmental Health Science and Research Bureau, Health Canada, 251 Sir Frederick Banting Driveway, Ottawa, ON K1A 0K9, Canada. E-mail: paul.white@canada.ca

Received 22 November 2019; Revised 27 November 2019; Accepted 30 November 2019

DOI: 10.1002/em.22351

Published online 3 December 2019 in Wiley Online Library (wileyonlinelibrary.com).

Environmental and Molecular Mutagenesis. DOI 10.1002/em

cohosh extract induces chromosome damage *in vitro* in human TK6 cells. Bhalli et al. (2013) showed that aristolochic acid, a constituent of *Aristolochia* spp., induces *Pig-a* locus mutations in the peripheral blood of orally treated rats. A recently published book provides an important overview of DNA-damaging agents that are naturally found in common plants (MacGregor 2019). In addition to plant constituents, numerous potent genotoxicants, including aflatoxin B1, acrylamide, heterocyclic amines such as PhIP (2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine), and polycyclic aromatic hydrocarbons (PAHs) such as benzo[*a*]pyrene, are introduced into common foods via cooking and microbial contamination (Bolger et al. 2010; Carthew et al. 2010; Benford et al. 2010a; Xu et al. 2014; Yebra-Pimentel et al. 2015). Lastly, numerous combustion-derived genotoxicants, including PAHs and related compounds (e.g., nitroarenes, oxy-PAHs, etc.), are unavoidably encountered via contact with environmental matrices such as indoor air, outdoor air, drinking water, and soil (Lundstedt et al. 2007; IARC 2010a; IARC 2010b; IARC 2014; IARC 2016; Long et al. 2017).

With respect to the notion that low-dose regions of genotoxicity DR functions are characterized by responses that do not exceed the ever-present background, the scientific literature is replete with examples; moreover, examples whereby lack of response at low doses is mechanistically understood (COM 2018; Jenkins et al. 2005; Johnson et al. 2009; Johnson et al. 2012; Zair et al. 2011; Thomas et al., 2013a; Guerard et al. 2015; Kirsch-Volders et al. 2000; Moreno et al. 2019; MacGregor et al. 2015a; Arimoto-Kobayashi et al. 1997; Lutz 2009; Pottenger et al. 2009). For example, in their study of the clastogenic effect of the alkylating agent ethylmethanesulfonate (EMS) in human AHH-1 cells, Zair et al. (2011) documented a shift of the zero-order kinetics portion of the concentration–response function to lower concentrations upon knockdown of the base excision repair enzyme *N*-methylpurine DNA glycosylase. Similarly, in their study of *N*-methyl-*N*-nitrosourea (MNU) mutagenicity in AHH-1 cells, Thomas et al. (2013a) revealed that the extent of the zero-order kinetics portion of the concentration–response relationship is dependent on the activity of O^6-methylguanine-DNA methyltransferase (MGMT). With respect to effects in bacteria, Arimoto-Kobayashi et al. (1997) noted that the zero-order kinetics portion of the mutagenicity concentration–response function in *Escherichia coli* exposed to *N*-nitrosodimethylamine (NDMA) is shifted to lower concentrations in alkyltransferase-deficient mutants (i.e., *ada* and *ogt* knockouts). With respect to *in vivo* DR functions, the numerous publications that followed the Viracept® contamination incident noted that the *in vivo* mutagenicity DR for EMS displays zero-order kinetics until the capacity to repair DNA adducts becomes saturated (Lutz 2009). In this latter case, the lack of effect at low doses was associated with alkyl group removal by MGMT (Daniels et al. 2004). Numerous researchers have also documented the involvement of translesion DNA polymerases in alterations of sensitivity to noteworthy mutagens. For example, Moreno et al. (2019) showed that the zero-order kinetics portion of an *in vitro* UVA DR function is reduced in polymerase η defective cells. Relatedly, Quinet et al. (2016) noted a similar outcome in UVC-exposed cells that are deficient in Xeroderma pigmentosum complementation group C (XPC), plus either polymerase Rev1 or Rev3L. Goullet de Rugy et al. (2016) noted that polymerase θ depleted cells are more sensitive to DNA double-strand breaks introduced by treatment with hydroxyurea. Importantly, the manifestation of zero-order kinetics at low doses can also result from phenomena such as detoxification, that is, the effect increases only after capacity is exceeded (MacGregor et al. 2015a). For an overview of mechanistic information used by authoritative organizations to infer nonlinear DR functions characterized by zero-order kinetics at low doses, the reader is referred to the International Workshop on Genotoxicity Testing reports by MacGregor et al., in addition to a recent statement by the UK Committee on Mutagenicity (COM) regarding the use of quantitative approaches for interpretation of genotoxicity DR data (MacGregor et al. 2015a, 2015b; COM 2018).

With respect to acceptance of genotoxicity as a *bona fide* toxicological endpoint for regulatory decision-making, it is well known that induced mutations in somatic and germ cells are associated with a wide range of human diseases; moreover, that mutation *per se* is an adverse health outcome (Heflich et al. 2019; Environmental Mutagen Society Committee 17 1975; Yauk et al. 2013, 2015). Initially, in the early to mid-20th century, germ cell mutation was understandably viewed as a prerequisite to heritable genetic disorders (Environmental Mutagen Society Committee 17 1975); indeed the Human Genome Mutation Database documents over a quarter of a million inherited human mutations, in over 10,000 genes, that are associated with genetic diseases (www.hgmd.cf.ac.uk/ac/index.php) (Stenson et al. 2017). Nevertheless, this recognition was eclipsed in the early 1970s by a focus on cancer, and the use of genotoxicity assessment to merely identify potential carcinogens (MacGregor et al. 2015a, 2015b; Heflich et al., 2019). Despite this focus, some regulatory authorities maintained that genotoxicity (i.e., mutation) must be viewed as a *bona fide* toxicological endpoint, and "notwithstanding the published numerical correlations between… mutagenic activity and the cancer bioassay, it is likely that an agent that induces mutations in the somatic cells of test animals will possess the potential to manifest this mutagenic activity as some adverse health effects" (DNHW 1993). Relatedly, numerous researchers have noted that somatic mutations manifested as tissue mosaicism have been associated with a wide range of metabolic, hematologic, and neuromuscular disorders (Youssoufian and Pyeritz 2002; Erickson 2010; Godschalk et al. 2020).

Environmental and Molecular Mutagenesis. DOI 10.1002/em

Although few researchers have documented the ability of xenobiotic exposures to induce somatic mosaicism in adults (Horibata et al. 2016; Dobrovolsky et al. 2011), several researchers have constructed adverse outcome pathways (AOPs) that conceptually connect DNA damage (i.e., key events) to chromosomal abnormality and mutation. These AOPs formally acknowledge that genotoxic effects such as chromosome damage and mutations are indeed *bona fide* toxicological endpoints (Yauk et al. 2013, 2015) (see aopwiki.org). For a detailed discussion of the issues pertaining to acknowledgement of mutation as a *bona fide* toxicological endpoint, as well as current efforts to improve the ability to assess mutagenic effects *in vivo*, and the interpretation of mutagenicity data in a risk assessment context, the reader is referred to the works by Heflich et al. (2020) and Godschalk et al. (2020).

Gradual acknowledgment that (1) genotoxicants are not rare, (2) genotoxicity DR functions display zero-order kinetics at low doses, and (3) genotoxic effects such as mutation are adverse effects in their own right, has contributed to worldwide interest in the development of quantitative methods for the analysis and interpretation of genotoxicity DR data; moreover, interest in the development of methods and approaches to interpret genotoxicity DR data for risk assessment and regulatory decision-making. Indeed, acknowledgment that the aforementioned assumptions are false necessitates the use of quantitative analyses to determine exposure levels associated with negligible risk of adverse human health effects (i.e., genetic damage). In this regard, it is important to acknowledge the influence of the aforementioned Viracept® contamination incident, which took place in 2007, on the development and use of methods underlying quantitative analysis of mutagenicity DR data for risk assessment (Lutz 2009; Muller et al. 2009; Muller and Gocke 2009).

IDENTIFYING A SUITABLE PoD METRIC FOR INTERPRETATION OF GENOTOXICITY DOSE–RESPONSE DATA

Although the development and establishment of standardized methods for quantitative interpretation of genotoxicity DR data can borrow from the well-established framework employed for interpretation of toxicological DR data in general, several impediments have stalled routine implementation of quantitative methods to interpret genotoxicity DR data in a regulatory context. This includes the development of standardized methods for quantitative analysis of genotoxicity DR data and determination of PoD metrics reflective of test article potency. Several researchers have evaluated various approaches for quantitative analysis of genetic toxicity DR data; moreover, evaluated DR PoD metrics such as the No Observed Genotoxic Effect Level (NOGEL), the Breakpoint Dose (BPD, also referred to as the Td or Threshold Dose), the Slope Transition Dose (STD), and the Benchmark Dose (BMD). More specifically, Gollapudi et al. (2013) analyzed EMS and methyl methanesulfonate (MMS) DR data, and critically evaluated the NOGEL, Td, and BMD. The follow-up work by Johnson et al. (2014) analyzed *N*-ethyl-*N*-nitrosourea (ENU) and MNU DR data, and critically evaluated the NOGEL, BPD, STD, and BMD. These works both noted that there are distinct advantages to the BMD approach, which employs computational algorithms to fit mathematical functions to DR data; more specifically, distinct advantages to an approach that mathematically interpolates the dose or concentration that corresponds to a small, predefined increase in response above background. Unlike the NOGEL, the BMD is not restricted to a study concentration/dose, and BMD confidence intervals can readily be calculated. Moreover, BMD does not suffer from the implicit impracticality of determining a dose that fails to elicit a response (i.e., the no effect level). In actual fact, the entire "no effect level" concept is fallacious and invalidated by the aphorism "Absence of Evidence is Not Evidence of Absence" (Altman and Bland 1995). Additionally, unlike the BPD and STD, the BMD is easily determined using readily available free software (e.g., BMDS and PROAST) and can be reliably estimated using standardized mathematical algorithms (Slob 2002; Slob and Setzer 2014). For a comprehensive overview of the advantages and disadvantages of the aforementioned PoD metrics, the reader is referred to the aforementioned work by Johnson et al. (2014).

The BMD approach, which was established in 1984 (Crump 1984), is broadly used for the interpretation of toxicological DR data, and it is now generally acknowledged that the BMD approach constitutes the most effective and flexible strategy for determining PoD metrics for comparative potency analysis, margin of exposure (MOE) determination, and, via extrapolation, determination of Health-based Guidance Values (HBGVs) that specify a human exposure limit (e.g., Permitted Daily Exposure or PDE). The predefined response increase employed for BMD determination (e.g., 10%) is known as the critical effect size (CES) or Benchmark Response (BMR), and the lower and upper confidence limits on the BMD for a given CES referred to as the BMDL and BMDU, respectively. Importantly, the ratio of the BMDL to the BMDU reflects the precision of the BMD and its potential utility for regulatory decision-making.

With respect to comparative potency analyses, several recent works have shown how the BMD combined covariate approach can be used to simultaneously analyze multiple DR functions and rigorously compare potency values (i.e., BMDs) across covariates such test article, sex, treatment duration, and cell type (Wills et al. 2016a; Wills et al. 2016b; Allemang et al. 2018; Guo et al. 2018). The results obtained can be used for compound potency ranking and

Environmental and Molecular Mutagenesis. DOI 10.1002/em

read across, to support mode of action determination, and to identify determinants of assay performance and utility (e.g., treatment duration, cell type, etc.). With respect to HBGV specification, the BMD approach can be employed to effectively define genotoxicity PoD values, which in turn can be adjusted to determine human exposure limits for risk assessment (i.e., HBGVs). Regarding BMD precision, it is important to emphasize that a large BMDL-to-BMDU ratio will limit the utility of the BMD for covariate analysis and HBGV determination. Regarding the former, large confidence limits will restrict that ability to detect significant differences across the covariate levels. With respect to the latter, a large BMDL-to-BMDU ratio will result in a large HBGV confidence limit. Although it is not possible to specify a BMDL-to-BMDU ratio to identify BMDs that are useful for regulatory decision-making, some authors have commented on the value of the ratio, and the repercussions of using imprecise BMD values. For example, Benford et al. (2010b) and Barlow et al. (2006) noted that the BMDL-to-BMDU ratio should be used as a tool for evaluating the statistical quality of the underlying DR data, and the overall utility of a BMDL as a reference PoD for regulatory decision-making. White et al. (2019) noted that large BMDL-to-BMDU ratios (e.g., >100) suggest that the DR data may not be suitable for regulatory use. Overall, prior to routine regulatory application of BMD-based approaches in genetic toxicology, it is necessary to emphasize the importance of routinely considering BMD precision; and moreover, to address several significant knowledge gaps. More specifically, although some researchers have endeavored to specify CES values for genotoxicity endpoints (Zeller et al. 2017), it will be necessary to more definitely define robust endpoint-specific CES values and determine appropriate values for the Assessment Factors (AFs) used to extrapolate below the BMD.

KNOWLEDGE GAPS IMPEDING ROUTINE USE OF THE BMD APPROACH FOR REGULATORY INTERPRETATION OF GENOTOXICITY DOSE-RESPONSE DATA

Lack of Endpoint-Specific CES Values

Regulatory organizations that support the use of the BMD approach also provide recommendations for CES. For example, the European Food Safety Authority (EFSA) recommends a CES of 5% for continuous toxicological endpoints such as organ weight, body weight and hematologic parameters, and 10% for quantal endpoints such as cancer (EFSA 2009; Hardy et al. 2017). Similarly, the US EPA recommends a CES of 10% for quantal endpoints such as cancer, and a CES of one standard deviation (SD) from the mean of the control (i.e., BMD_{1SD}) for continuous endpoints (USEPA 2012; Wignall et al. 2014; Haber et al. 2018). The 1SD approach has been criticized, particularly for endpoints with low response variability whereby it is unlikely that a 1SD change from control (i.e., background) could be deemed adverse (Haber et al. 2018). Conversely, for endpoints with high control variability, the 1SD approach will yield larger CES values, that is, the percentage increase corresponding to a 1SD increase above control will be relatively large. Larger CES values will yield larger BMD and BMDL values, which may be less desirable from a regulatory point of view (i.e., less restrictive). Nevertheless, the USEPA recommends that a BMR corresponding to 1SD increase over control be presented for comparative purposes; the value might be used when there is no consensus on a biologically defensible BMR for the endpoint under consideration (USEPA 2012).

There is considerable inconsistency with respect to the CES values used for continuous endpoints; moreover, the rationale used to select an appropriate CES (Haber et al. 2018). For example, whereas the EFSA recommends 5%, a value that is based on detailed cross-study analyses of No Observed Adverse Effect Level (NOAEL) values for a variety of toxicological endpoints (EFSA 2009; Hardy et al. 2017), the USEPA notes that low CES values may necessitate undesirable extrapolation outside the range of observations (USEPA 2012). The USEPA further notes that acceptance of CES values lower than 10% should be based on biological considerations and the range of observed responses (USEPA 2012). EFSA similarly notes that "the default BMR (CES) may be modified based on statistical or biological considerations" (Hardy et al. 2017), for example, large confidence intervals for BMDs associated with low CES values. For a complete discussion regarding the use of the BMD approach in general, and considerations for selection of an appropriate CES, the reader is referred to the recent work of Haber et al. (2018), as well as the aforementioned works by Hardy et al. (2017), the EFSA (2009), and the USEPA (2012).

There is an acute paucity of serious discussions regarding appropriate CES values for genotoxicity endpoints such as mutation and chromosome damage. Some researchers have employed a CES of 10% for analysis of *in vivo* mutagenicity DR data (Johnson et al. 2014); others have noted that 10% is generally too low for genotoxicity endpoints (Zeller et al. 2017). More specifically, the recent work of Zeller et al. (2017), which acknowledged the need to address CES determination, proposed delineation of CES values for genotoxicity endpoints using strategies employed to scrutinize historical background response values and define a statistically meaningful positive response. Their work noted that CES values for *in vivo* genotoxicity endpoints such as chromosome damage, mutation, and DNA strand breaks should be in the range of 34–76%. Nevertheless, since their work is based on analysis of historical controls, and does not reflect endpoint-specific dynamic range (i.e., fold-change maxima); the values should be regarded as minima. Importantly, the work of Slob (2016) clearly demonstrated that robust estimation of CES; moreover,

Environmental and Molecular Mutagenesis. DOI 10.1002/em

CES values that reflect what the author denotes as a small adverse effect, requires careful consideration of DR shape and related variations in DR function parameters across endpoints and compounds (Slob 2016). More specifically, Slob presents an effect size (ES) theory that demonstrates how CES determination can be based on *M*, the endpoint-specific maximum, and/or its surrogate *s*, the within-group standard deviation. For example, the results presented by Slob, which evaluated DR data for 27 endpoints, provide *s* confidence limits indicating that the CES corresponding to a small change in *in vivo* chromosome damage incidence (i.e., micronuclei or MN frequency) should be in the range of 71–79%. Similarly, although based on more limited data, the CES for *Pig-a* mutation frequency in peripheral blood should be in the range of 58–70%, and the CES for transgene mutation frequency in BigBlue® rat liver should be in the range of 46–74% (Slob 2016). This latter value is consistent with the values noted by Wills et al. (2017). In their analyses of Transgenic Rodent (TGR) mutagenicity data, Wills et al. (2017) noted that CES values for *in vivo* mutagenicity should be in the range of 18–66%. Although the approach advocated by Slob can certainly be used for robust determination of endpoint-specific CES values, it requires collection, collation, and analysis of large data sets containing DR data across a wide range of studies and test articles. To date, such detailed analyses have not been conducted, and there is no general consensus regarding appropriate CES values for genotoxicity endpoints. Moreover, there is debate about how to define a unique genotoxicity endpoint. More specifically, it is not clear whether the induction of mutations in TGRs should be considered an endpoint, or whether induction of mutations in each TGR tissue should be considered unique endpoints. Work to critically consider CES options are ongoing and endpoint-specific CES values will be more definitely specified in a forthcoming work.

Lack of a Standardized Framework for Regulatory Interpretation of Genotoxicity BMD Values

In addition to the use of the BMD approach to compare genotoxic potency values (i.e., BMDs) across covariates such as compound, sex, tissue, tissue sampling time, and cell type, BMD values can also be used for risk assessment. Importantly, although the choice of CES has a negligible effect on potency ranking across covariate levels (Wills et al. 2016a), the choice of CES is critically important for determination of BMD values to be used for risk assessment and regulatory decision-making. More specifically, since CES specification will impact the value of the BMD and its confidence limits, it will necessarily influence the value of metrics employed for risk assessment, for example, MOEs and HBGVs. Therefore, there is an acute need to determine CES values for genotoxicity endpoints such as mutation and chromosome damage; more specifically, a need to specify CES values that can be used to determine BMDs, which can, in turn, be confidently employed for human health risk assessment.

The MOE approach simply examines the ratio between the lower confidence limit of the BMD (i.e., the BMDL) and the estimated human exposure. This ratio, which as the name implies is the effective margin of safety between the dose that elicits a set fractional increase in response and the estimated human exposure, is officially endorsed by the EFSA (EFSA 2009; Hardy et al. 2017). It has been frequently employed to assess the risks posed by unavoidable contaminants in food (Bolger et al. 2010; Carthew et al. 2010; Coulet et al. 2010; Renwick et al. 2010; Smith et al. 2010; Williams et al. 2010; Benford et al. 2010a, 2010b).

HBGVs are human exposure limit values such as the PDE, the tolerable daily intake (TDI), acceptable daily intake, or reference dose (IPCS 2014; ICH 2016) (ICH is defined as International Council on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use). Calculations of deterministic HBGVs, such as a PDE for a pharmaceutical impurity or a TDI for an unavoidable environmental contaminant, can be calculated as the quotient of the body weight-adjusted BMDLxx (e.g., BMDL for a specified CES in mg per individual) and the product of several AFs (assessment factors).

An AF, which is a general term that covers a range of metrics alternatively referred to as an uncertainty factor, an extrapolation factor, a safety factor, or an adjustment factor, generally accounts for uncertainties related to interspecies extrapolation, variability in human sensitivity (i.e., intraspecies extrapolation), study duration, and effect severity (Vermeire et al. 1999; ECHA 2012). Additional factors are sometimes used to account for the route of exposure, exposures of children, suspected interactions in mixtures, database insufficiency, scientific uncertainty, or expert judgment (Vermeire et al. 1999; Ritter et al. 2007; ECHA 2012). The literature pertaining to the AF values employed to calculate human exposure limits is unfortunately inconsistent and convoluted, with different values recommended by different regulatory authorities for different types of substances, and different regulatory domains (i.e., food, drinking water, soil, pharmaceutical impurities, etc.). Consequently, this work only provides a brief overview of the values employed in different regulatory scenarios. For a detailed discussion regarding the values commonly used for chemical risk assessment (i.e., to calculate HBGVs), and their origins and rationale, the reader is referred to several documents published by the European Chemicals Agency (ECHA), the EFSA, the USEPA, the ICH, the United States Food and Drug Administration (USFDA), and the WHO (IPCS 1994, 1999; WHO 2001; USFDA 2005; WHO 2005; EFSA 2009; ECHA 2012, 2014; IPCS 2014; USEPA 2014; ICH 2016; Hardy et al. 2017).

Environmental and Molecular Mutagenesis. DOI 10.1002/em

ADDRESSING THE KNOWLEDGE GAPS IMPEDING ROUTINE USE OF GENOTOXICITY DOSE-RESPONSE DATA FOR RISK ASSESSMENT AND REGULATORY DECISION-MAKING

Establishing CES Values for Genetic Toxicity Endpoints

Since the lack of endpoint-specific CES values is restricting the use of genetic toxicity DR data for risk assessment, it is essential to determine appropriate CES values for genotoxicity endpoints such as mutation and chromosome damage. The aforementioned work of Slob employs the ES theory described therein to calculate endpoint-specific CES values for a wide range of endpoints including chromosome damage in rodents (i.e., micronuclei in peripheral blood) (Slob 2016). Thus, a laudable objective of future work is the use of the Slob ES theory to determine appropriate CES values for other genotoxicity endpoints. The ES theory demonstrates that harmonized, endpoint-specific CES values can be obtained by scaling ES values, expressed as a percent change in mean response, to the logarithm of M (maximum ES). Since M is often difficult to estimate, s (i.e., the within-group standard deviation) would be determined as a proxy, and s values then used to mathematically determine endpoint-specific CES values. The complexity of the available DR data for endpoints such as mutation in TGR tissues, which includes a wide range of experimental protocols and dosing schedules, tissues and tissue sampling times, and exposure routes, will necessitate detailed analysis of s. More specifically, prior to determining a CES value for a particular genotoxicity endpoint, it will be necessary to determine the dependency of s on experimental parameters such as exposure route, dosing schedule, tissue, and tissue sampling time.

More recent analyses of TGR and *Pig-a* mutagenicity DR data, some of which are hitherto unpublished, indicate that endpoint-specific CES values should be approximately 60% and 56%, respectively. Although these values are based on a detailed analysis of numerous data sets that cover a broad range of compounds, TGR strains and loci, routes of administration and dosing schedules, and tissues, the values should be treated as preliminary. Determination of robust, definitive CES values for genetic toxicity endpoints will require a concerted effort to investigate the dependence of endpoint-specific values of M and/or s on the aforementioned covariates. Fortunately, since extensive databases of TGR and *Pig-a* DR data are now available, the work can be readily conducted. More specifically, analyses to determine CES values can use the extensive Transgenic Rodent Assay Information Database (TRAiD) currently housed at Health Canada in Ottawa, and the extensive *Pig-a* assay database currently housed on a University of Maryland server (www.pharmacy.umaryland.edu/centers/cersi-files) (Shemansky et al. 2019).

Establishing Data-Derived AFs for Regulatory Interpretation of Genotoxicity BMD Values

As noted earlier, the determination of deterministic HBGVs such as PDEs or TDIs (i.e., human exposure limits) can be calculated as the quotient of the body weight-adjusted BMDL and the product of several AFs. Deterministic HBGVs, which specify the exposure level below which the risk of adverse human health effects is deemed negligible, are generally calculated using four AFs to account for (1) interspecies differences (e.g., experimental animal to human), (2) intraspecies variability (i.e., variability amongst humans), (3) study duration, and (4) effect severity. For example, an HBGV such as PDE or TDI can be calculated as indicated below.

$$\text{HBGV} = \frac{\text{BMDL x weight adjustment (e.g., kg per individual)}}{\text{AF1} \times \text{AF2} \times \text{AF3} \times \text{AF4}}$$

As noted, additional AFs can be employed when the database of available information is deemed to be insufficient for effective risk assessment (Ritter et al. 2007; Dankovic et al. 2015; ICH 2016). For example, additional AFs can compensate for suspected interactions of chemicals in mixtures or likely exposures of children (Ritter et al. 2007). The following sections provide a brief overview of commonly used AF values, a summary of preliminary work to determine AF values for risk assessment of genotoxic substances, and a brief outline of opportunities for further research.

The interspecies AFs account for toxicokinetic (TK) and toxicodynamic (TD) differences between experimental animals (e.g., rat, mouse, dog, etc.) and humans. Although some regulatory guidance documents (e.g., WHO/IPCS) recommend separate accounting for TK and TD differences (IPCS 1999; WHO 2001; WHO 2005), calculations often only include an allometrically defined TK adjustment based on body size (USFDA 2005; IPCS 2014; USEPA 2014; Dankovic et al. 2015; ICH 2016). These values, which generally follow USFDA recommendations, are 12 for mouse, 5 for rat, 2.5 for rabbit, 2 for dog, and 3 for monkey. (ICH 2016). Although the empirical basis for the value is unclear, the AF employed to account for interspecies TD differences is often 3.16 (i.e., $10^{0.5}$) (IPCS 1999; WHO 2001, 2005).

The capricious AF used to account for intraspecies variability (i.e., variability in human sensitivity) is generally 10 (IPCS 1994; WHO 2005; ECHA 2012; ICH 2016), and, in the absence of "appropriate data" (Ritter et al. 2007), most regulatory authorities default to an interspecies and intraspecies AF of 100, that is, 10 for each. Although the WHO has indicated that this default value may not be ideal or "correct," it reflects the common approach used to derive HBGVs for the general population (WHO 2005).

Environmental and Molecular Mutagenesis. DOI 10.1002/em

Several regulatory authorities have noted that generation and interpretation of appropriate data can contribute to further improvement of the scientific basis for commonly used AFs. Indeed, robust AFs are particularly important if HBGVs are to be based on a robust PoD metric, for example, BMD and its confidence limits. Improved AF values, which can be based on careful consideration of available mechanistic information (e.g., comparative interspecies and intraspecies TK and TD data), have been referred to as data-derived extrapolation factors (DDEFs) or chemical-specific adjustment factors (Haber et al. 2001; WHO 2001; IPCS 2014; USEPA 2014). Interestingly, the USEPA has noted that ratios of toxicologically relevant metrics (e.g., AUC, BMD, NOAEL, etc.) can be used to derive empirical interspecies and intraspecies AF values (i.e., DDEFs) (Vermeire et al. 1999; USEPA 2014). More specifically, a human-to-animal ratio of toxicologically relevant metrics can be used to empirically determine interspecies AF values; intraspecies values can be derived using ratios of metrics for sensitive humans to metrics for humans that are representative of the general population. It should be noted that intraspecies AFs to account for differences in human sensitivity often vary according to the subpopulation under consideration. For example, although regulatory authorities often recommend 10 for the general population, some recommend lower values (e.g., 2–6) for occupational settings (i.e., workers) (Li and Sung 1999; ECHA 2012; Dankovic et al. 2015).

With respect to study duration, different regulatory authorities recommend different AFs. For example, the ECHA recommends 6 for subacute (i.e., 28 day) to chronic (i.e., 1.5–2 years) and 2 for sub-chronic (i.e., 90 days) to chronic (ECHA 2012, 2014). In contrast, the ICH recommends 1 for half-lifetime studies (i.e., 1 year for rodents), 5 for 90-day studies, and 10 for all study durations less 90 days (ICH 2016). An overview presented in Vermeire et al. (1999) indicates that regulatory authorities employ deterministic subacute to sub-chronic values from 3 to 100, and sub-chronic to chronic values ranging from 1 to 10. In some cases, a more complex approach is employed to account for age-related response variability (Calabrese and Gilbert 1993). Overall, regulatory authorities employ a range of deterministic values to account for less-than-chronic exposure; a detailed overview is beyond the scope of this work (ECHA 2012).

To date, there has not been any concerted effort to determine AF values to support the interpretation of genotoxicity PoD values in a regulatory context, and there is no consensus regarding which, if any, published AF values are appropriate for genotoxicity endpoints. Nevertheless, the aforementioned work by Johnson et al. (2014) used the BMDL values associated with *in vivo* mutagenicity DR data to establish human exposure limits for ENU and MNU (Johnson et al. 2014). In this particular case, in the absence of any recommendations for interpretation of genotoxicity PoD values, the authors employed allometrically determined AF values to account for interspecies differences (i.e., 12 for mouse and 5 for rat), and 10 to account for intraspecies differences in human sensitivity. In addition, they employed an additional AF of 10 to account for interspecies TD differences. They did not employ AF values to account for study duration or effect severity. Muller and Gocke (2009) used the same interspecies AF for TK, and the same intraspecies AF for human sensitivity in their determination of a PDE for EMS (Muller and Gocke 2009). They also included maximal AF values of 10 for each of treatment duration and effect severity.

The AF values employed for the aforementioned regulatory evaluations of genotoxicity PoDs cannot be rigorously defended; thus, there is an acute need to employ an empirical approach to determine appropriate AF values for genotoxicity endpoints. Such an approach, which would be consistent with the USEPA's recommendation to use mechanistic information on receptor affinity, enzyme inhibition, and/or repair of DNA or tissue damage for determination of DDEFs (USEPA 2014), could provide robust, pragmatic AF values for genotoxicity risk assessment and concomitant regulatory decision-making. From an operational point of view, a 2014 USEPA guidance document notes that, in the absence of sufficient data to develop robust TK and/or TD models, it is not necessary to employ single, deterministic default values to account for risk assessment uncertainty. As an alternative, the document notes that DDEF values can be calculated as ratios of PoD values from key studies. Such an approach has also been advocated by Vermeire et al. (1999) who indicate that NOAEL ratios, or distributions of NOAEL ratios, can be used to specify appropriate AFs to account for uncertainties related to interspecies and intraspecies extrapolation and study duration. In this regard, the peer-reviewed scientific literature can be employed to identify relevant PoD values, and ratios or ratio distributions used to determine AF values that are empirically supported by appropriate mechanistic considerations. As an example, published data on differential ultraviolet light sensitivity of cultured lymphoid cells from normal humans, and cultured lymphoid cells from humans with XPC mutations (i.e., nucleotide excision repair [NER] deficiency), could be used to determine an appropriate AF value to account for interindividual variation in human sensitivity (Cleaver et al. 2007). Using data from Cleaver et al., such a comparison indicates a differential sensitivity of 5.5- to 12.6-fold.

In an effort to determine an intraspecies AF value to account for variability in human sensitivity, the PoD ratio approach was employed herein to investigate the differential genotoxicant sensitivity of cells/animals deficient in a variety of DNA repair pathways (e.g., alkylation removal, base excision repair [BER], NER, strand break repair, and translesion synthesis [TLS]). For example, a study by Moreno et al. (2019) documented the differential sensitivity

Environmental and Molecular Mutagenesis. DOI 10.1002/em

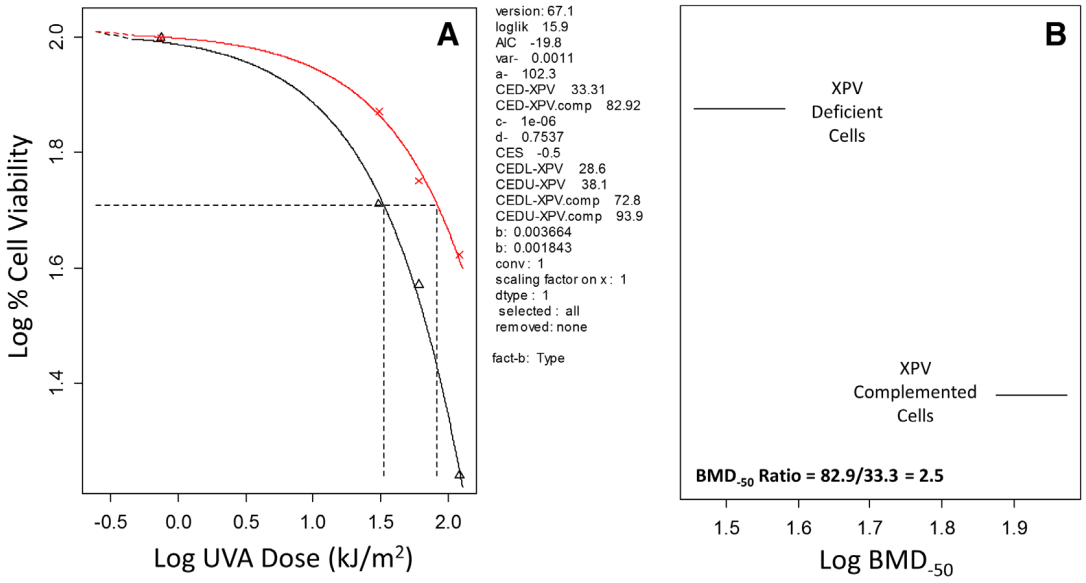

Fig. 1. The effect of TLS deficiency on the survival of UVA-treated cells. (**A**) Dose–response relationship showing the effect of UVA treatment on the survival of human fibroblasts. Red plotting symbols and fitted function show the results for XPV-deficient cells (i.e., Pol η deficient) complemented with the wild-type XPV gene (i.e., POLH). Black shows the results for XPV-deficient cells (i.e., Pol η deficient). Vertical dotted lines extending to the horizontal axis indicate the UVA dose required to reduce cell viability by 50% (i.e., BMD_{-50}). (**B**) Comparison of BMD values for the two cell types investigated (i.e., XPV-deficient vs. XPV-complemented). The left and right extremes of the horizontal lines indicate the lower and upper limits of the BMD 90% confidence interval, respectively. Dose–response data were analyzed using PROAST v67.7 with cell type as a covariate. Data were obtained from Moreno et al. (2019). Data presented in figures were extracted using WebPlotDigitizer (https://apps.automeris.io/wpd/).

of Pol η (Pol eta) deficient human fibroblasts (i.e., TLS deficient cells) exposed to ultraviolet light (i.e., UVA). The differential sensitivity is consistent with bypass of UVA-induced DNA damage (i.e., cyclobutane dimers) that elicits replication fork stalling and *S*-phase arrest (Moreno et al. 2019); comparative analysis yields a BMD ratio of 2.5 (Fig. 1). Similarly, a study by Allan et al. (1998) documented the differential sensitivity of murine embryonic fibroblasts deficient in *N*-methylpurine DNA glycosylase, (i.e., BER deficient cells) exposed to the potent mutagen *bis*-chloroethylnitrosourea. The differential sensitivity reflects the importance of DNA alkylation repair; comparative BMD analysis yields a ratio of 2.1 (Fig. 2). Table I summarizes the results of the comparative PoD analyses conducted to date. The results presented indicate differential sensitivity values, expressed as PoD ratios, of 4.0–15.8, 1.6–20.3, 5.5–12.6, 1.4–2.6, and 2.1–5.5 for alkylation removal, BER, NER, strand break repair, and TLS, respectively. Thus, the default deterministic intraspecies AF employed by numerous regulatory authorities (i.e., 10) appears reasonable for genotoxicity endpoints. In fact, it may be too conservative. Since the results summarized in Table I are based on review of data from only 11 studies; moreover, on a preponderance of studies on animal cells, the results should be regarded as preliminary. Going forward, it will be necessary to conduct a comprehensive analysis of the scientific literature to further scrutinize the impact of deficiency in compensatory responses (e.g., DNA repair capacity, metabolism, etc.) on the sensitivity of cultured cells and/or experimental animals to a range of genotoxicants. The range, distribution, and central tendencies of comparative sensitivity values will permit determination of an appropriate, robust genotoxicity AF value to account for intraspecies variability, that is, variability in human sensitivity. Moreover, the analysis will permit critical evaluation of the default deterministic values employed by regulatory authorities for other toxicological endpoints (e.g., cancer). More specifically, a critical evaluation of the general population 10-fold intraspecies AF advocated by numerous regulatory authorities (Vermeire et al. 1999; WHO 2001; Ritter et al. 2007; ECHA 2012; Dankovic et al. 2015). Importantly, such an approach is consistent with the ECHA statement that "the probability of genetic alterations [in humans]…may be dependent on the efficiency of DNA repair and cell cycle control" (ECHA 2012).

There is no consensus on the appropriate treatment/study duration AF for *in vivo* genotoxicity endpoints such as mutation and chromosome damage. Regulatory authorities such as the ICH and the ECHA recommend study duration AFs of 6–10 to compensate for studies where the treatment duration cannot be classified as chronic (i.e., >half-lifetime). Since

Environmental and Molecular Mutagenesis. DOI 10.1002/em

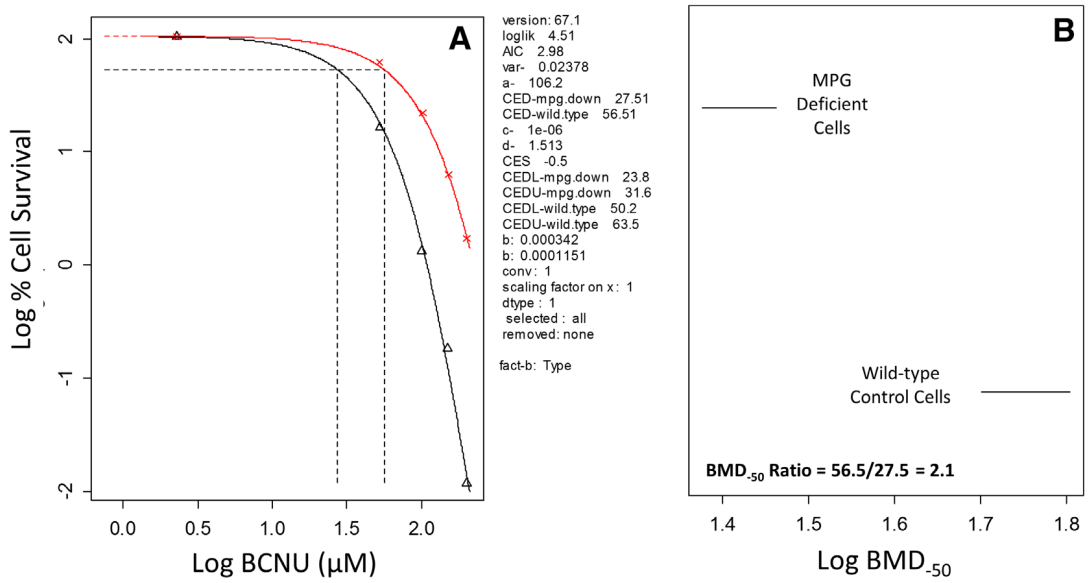

Fig. 2. The effect of BER deficiency on the survival of *bis*-chloroethylnitrosourea (BCNU)-treated cells. (**A**) Dose–response relationship showing the effect of BCNU treatment on the survival of murine ES (embryonic stem cells). Red plotting symbols and fitted function show the results for wild-type clones. Black shows the results for MPG-deficient clones (i.e., *N*-methylpurine DNA glycosylase deficient). Vertical dotted lines extending to the horizontal axis indicate the BCNU concentration required to reduce cell viability by 50% (i.e., BMD_{-50}). (**B**) Comparison of BMD values for the two cell types investigated (i.e., MPG-deficient vs. wild-type clones). The left and right extremes of the horizontal lines indicate the lower and upper limits of the BMD 90% confidence interval, respectively. Dose–response data were analyzed using PROAST v67.7 with cell type as a covariate. Data were obtained from Allan et al. (1998). Data presented in figures were extracted using WebPlotDigitizer (https://apps.automeris.io/wpd/).

OECD guidelines for endpoints such as somatic mutations in TGRs (e.g., MutaMouse, BigBlue rat and mouse, *gpt* delta rat and mouse) recommend treatment duration of 28 days (OECD 2013), it is tempting to employ an AF of 1 for TGR studies that treat animals for a minimum of 28 days. A similar argument could be made for the peripheral blood *Pig-a* gene mutation assay (Gollapudi et al. 2015). However, it is not clear if such a paradigm can be justified; moreover, it is not clear what AF values should be employed for treatment durations less than 28 days. In the aforementioned PDE determination for EMS by Muller and Gocke, the authors argued that a factor of 10 for a 30-day repeat-dose treatment is a worst-case value; moreover "that the NOEL is not expected to decrease if the treatment duration would be extended" (Muller and Gocke 2009). In addition, the aforementioned work of Johnson et al. (2014), which calculated human exposure limits for ENU and MNU, did not employ an AF for study duration. Interestingly, the DR data used by Johnson et al. to calculate PoDs for MNU and ENU were based on a 28-day repeat-dose regimen and a single acute treatment, respectively.

Interestingly, comparative analysis of published data for *Pig-a* mutations in red blood cells (RBCs) of rats orally treated with aristolochic acid indicates, for matched sampling times (i.e., matched manifestation times), an 11.7-fold difference between the BMDs for 3-day and 28-day repeat-dose treatment regimens (Fig. 3) (Bhalli et al. 2013). Similarly, comparative analysis of data for *Pig-a* mutations in RBCs of rats orally treated with MMS (Dertinger et al. 2012) indicates that, for matched post-exposure sampling times, the ratio of the 3-day BMD to the 28-day BMD is 5.2 (Fig. 4). A preliminary comparative overview of matched *Pig-a* mutation studies with both 3- and 28-day repeat-dose treatment regimens indicates, for 7 compounds, a BMD ratio of 4.6–14.8, that is, 28-day repeat-dose treatment DR data yield test article potency values that are 4.6- to 14.8-fold lower than those determined using 3-day repeat-dose treatment DR data (Bhalli et al. 2013; Torous 2012; Dertinger et al. 2011, 2012, 2014). The maximum ratio, which is aligned with several of the deterministic study duration AF values mentioned earlier, will yield lower human exposure limit values. Interestingly, the DR data analyzed to date (i.e., across 7 compounds) indicate 5th and 95th percentiles of the 3- to 28-day BMD ratios of 4.7 and 13.5, respectively. Thus, the analyses seem to suggest that, to account for treatment durations less than 28 days, a deterministic AF of 10 may be appropriate. However, most regulatory authorities would categorize a 28 day treatment as subacute relative to half-lifetime (i.e., 1 year for rodents), and it is not clear whether a 28-day repeat-dose treatment can be regarded as sufficient (i.e., chronic) for the endpoint. More specifically, at the present time, it is

TABLE I. Differential Sensitivity of Cells Deficient in Various DNA Repair or Lesion Bypass Pathways

Repair enzyme	Repair pathway	Endpoint	Exposure conditions	Endpoint effect ratio	Reference
(MPG) *N*-methylpurine DNA glycosylase[a]	Base excision repair	Frequency of micronucleated binucleate cells	AHH-1 cells, ethyl methanesulfonate	2.1[b]	Zair et al. (2011)
(MPG) *N*-methylpurine DNA glycosylase	Base excision repair	Cell survival	Mouse ES cells, methyl methanesulfonate	1.6[b]	Allan et al. (1998)
(MPG) *N*-methylpurine DNA glycosylase	Base excision repair	Cell survival	Mouse ES cells, *bis*-chloronitrosourea	2.1[b]	Allan et al. (1998)
(MPG) *N*-methylpurine DNA glycosylase	Base excision repair	Cell survival	Mouse ES cells, mitomycin C	1.9[b]	Allan et al. (1998)
(MPG) *N*-methylpurine DNA glycosylase	Base excision repair	Cell survival	Mouse ES cells, $MeOSO_2(CH_2)_2$-Lex	2.4[b]	Engelward et al. (1996)
(MPG) *N*-methylpurine DNA glycosylase	Base excision repair	Frequency of chromosome breaks	Mouse ES cells, Me-Lex	8.5[b]	Engelward et al. (1998)
(MPG) *N*-methylpurine DNA glycosylase	Base excision repair	Frequency of chromosome breaks	Mouse ES cells, methyl methanesulfonate	20.3[b]	Engelward et al. (1998)
XPC (Xeroderma pigmentosum complementation group C)	Nucleotide excision repair	Cell survival	Primary human lymphoid cells, UV (254 nm) exposure	5.5 to 12.6[c]	Cleaver et al. (2007)
MGMT (O^6-methylguanine-DNA methyltransferase	Alkylation removal	Frequency of mutations at *hprt* locus	AHH-1 cells, *N*-methyl-*N*-nitrosourea	5.5[d]	Thomas et al. (2013a)
MGMT (O^6-methylguanine-DNA methyltransferase	Alkylation removal	Dermal papilloma yield	Topical application of *N*-methyl-*N*-nitrosourea	15.6[e]	Becker et al. (2014)
MGMT (O^6-methylguanine-DNA methyltransferase	Alkylation removal	Cell survival	Bone marrow cells from *Mgmt* null mouse, methyl nitrososurea *in vitro*	4.0[b]	Glassner et al. (1999)
MGMT (O^6-methylguanine-DNA methyltransferase	Alkylation removal	Cell survival	Bone marrow cells from *Mgmt* null mouse, temozolomide *in vitro*	15.8[b]	Glassner et al. (1999)
Polymerase η	TLS (translesion synthesis)	Cell survival	XPV and XPV-complemented cells, UVA exposure	2.4[b] to 4.5[f]	Moreno et al. (2019)
Polymerase η	TLS (translesion synthesis)	Frequency of DNA double-strand breaks, as γH2AX foci	XPV and XPV-complemented cells, UVA exposure	4.7[f]	Moreno et al. (2019)
Polymerase η	TLS (translesion synthesis)	Frequency of DNA strand breaks, alkaline comet assay	XPV and XPV-complemented cells, UVA exposure	5.5[f]	Moreno et al. (2019)
Polymerase η	TLS (translesion synthesis)	Cell survival	Control cells and Pol η knock-down, UVC exposure	2.3[b]	Rey et al. (2009)
Polymerase η	TLS (translesion synthesis)	Frequency of aberrant metaphases	Control cells and Pol η knock-down, spontaneous	2.2 to 3.0[f]	Rey et al. (2009)
Polymerase η	TLS (translesion synthesis)	Frequency of aberrant metaphases	XPV and Pol η supplemented, spontaneous	3.4[f]	Rey et al. (2009)
Polymerase η	TLS (translesion synthesis)	Frequency of DNA double-strand breaks, as γH2AX foci	Control cells and Pol η knock-down, spontaneous	2.1[f]	Rey et al. (2009)
Polymerase θ	Strand break repair	Cell survival	Control cells and Pol θ knock-down, Ara-C exposure	2.0[b]	Goullet de Rugy et al. (2016)
Polymerase θ	Strand break repair	Cell survival	Control cells and Pol θ knock-down, hydroxyurea exposure	1.4[b]	Goullet de Rugy et al. (2016)

(Continues)

Environmental and Molecular Mutagenesis. DOI 10.1002/em

TABLE I. (Continued)

Repair enzyme	Repair pathway	Endpoint	Exposure conditions	Endpoint effect ratio	Reference
Polymerase θ	Strand break repair	Cell survival	Control cells and Pol θ over-expression, hydroxyurea exposure	2.6[b]	Goullet de Rugy et al. (2016)
Polymerase θ	Strand break repair	Frequency of DNA strand breaks, alkaline comet assay	Control cells and Pol θ knock-down, hydroxyurea exposure	2.1[f]	Goullet de Rugy et al. (2016)

[a]shRNA knock-down; MPG protein levels 27% of control.
[b]Based on BMD analysis with repair capacity as covariate, that is, ratio of BMD_{100} values for increasing responses, or ratio of BMD_{-50} values for decreasing responses.
[c]Ratio of D_{50} values, the dose associated with 50% cell survival (Cleaver et al. 2007, Table I).
[d]Maximum mutation frequency ratio (Thomas et al., 2013a, Fig. 5).
[e]Based on BMD analysis with repair capacity as covariate, repair over-expression vs. wild-type (Becker et al., 2014).
[f]Based on comparisons of responses for a single dose (e.g., Moreno et al. 2019, Fig. 4a).

not known whether, and to what extent, mutagenicity BMD values for endpoints such as *Pig-a* mutations in rat RBCs will decline for repeat-dose treatment durations longer than 28 days. Thus, it seems imperative that future research investigate and compare mutagenicity BMD values for repeat-dose treatment durations extending from 28 days to at least half-lifetime. More specifically, future research will need to investigate the empirical relationship between rodent BMDs and repeat-dose treatment duration, for durations extending from acute to half-lifetime (i.e., 1 year). Such investigations would determine the extent to which BMD declines with increasing treatment duration; moreover, whether the BMD reaches a steady-state minimum. The dose-rate corresponding to that minimum would be the repeat-dose treatment duration required to set the study duration AF to 1.

With respect to effect severity and reversibility, some regulatory authorities recommend use of 10 as a default AF for endpoints such as neurotoxicity, carcinogenicity, fetal toxicity, and teratogenicity (WHO 2001, 2005; Ritter et al. 2007; ICH 2016); there is currently no accepted value for genotoxic effects. The severity of genotoxic effects will clearly be dependent on the type of effect (e.g., DNA lesions, DNA strand break, mutations, chromosome damage), the consequence of the genetic alteration (e.g., change in protein-coding sequence), the location of the alteration in the genome (e.g., exon vs. intron), and the type of cell affected (i.e., germ cell vs. somatic). Some types of damage can be readily repaired or bypassed (e.g., strand breaks and bulky adducts); the severity and consequences of this type of damage might be considered relatively benign. That being said, lesion bypass can be error-free or error-prone. Nevertheless, because of the empirical and mechanistic connections with disease, mutations and/or chromosome damage might be regarded as far more severe (Youssoufian and Pyeritz 2002; Erickson 2010; Bonassi et al. 2011; Yauk et al. 2015; Marchetti et al. 2016; Meier et al. 2017; Stenson et al. 2017; Godschalk et al. 2020; Tomasetti 2019; Zhang et al. 2019). The location of mutations and/or chromosome damage within the genome will affect the consequences of the event, and relatedly, the risk of adverse health effect. Unfortunately, since TGRs commonly used to assess treatment-induced mutations can only examine mutations at transgenic loci, it is currently difficult to assess mutations elsewhere in the genome. Although recent developments now permit the use of Next Generation Sequencing technology to enumerate rare mutations throughout the genome of any organism (i.e., error-corrected sequencing) (Salk et al. 2018), the technology is relatively new and sample scoring is prohibitively expensive. Ability to score and scrutinize treatment-induced mutations throughout the genomes of exposed organisms, including humans, is certainly a promising area for further research. A more detailed

Environmental and Molecular Mutagenesis. DOI 10.1002/em

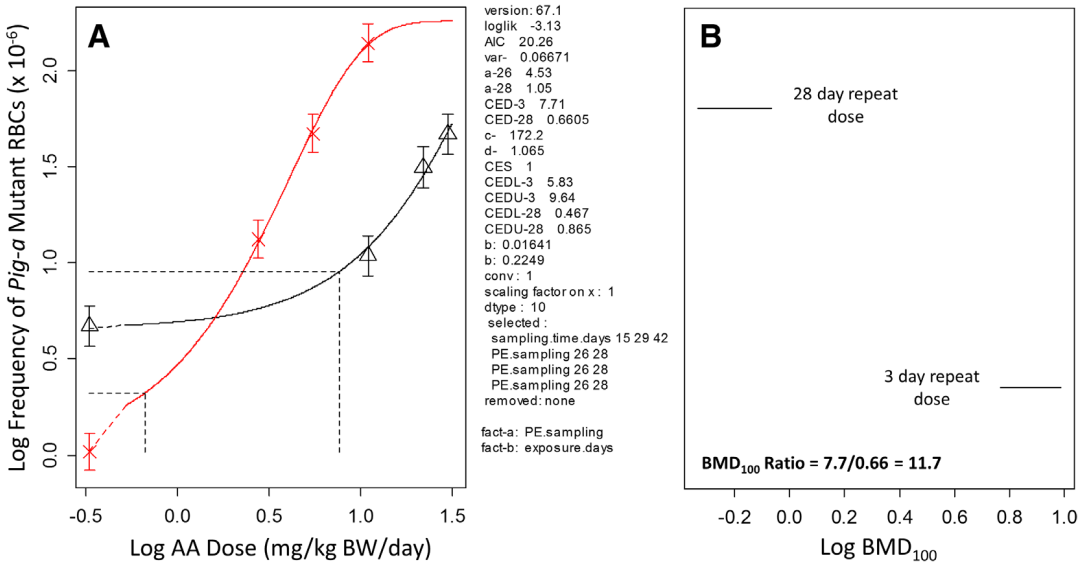

Fig. 3. The effect of treatment duration on the frequency of *Pig-a* mutant cells (i.e., CD59$^-$ cells). (A) Dose response relationship showing the frequency of *Pig-a* mutant rat RBCs (red blood cells) vs. oral dose of AA (aristolochic acid). Red plotting symbols and fitted function show the results for the 28-day repeat-dose treatment, black shows the results for the 3-day repeat-dose treatment; post-exposure sampling times are 26 and 28 days, respectively. Vertical dotted lines extending to the horizontal axis indicate BMD_{100} values (i.e., doubling dose). Plotting symbols show the mean values for each dose, error bars show the standard deviation. (**B**) Comparison of BMD_{100} values for the two treatment durations investigated. The left and right extremes of the horizontal lines indicate the lower and upper limits of the BMD 90% confidence interval, respectively. Dose–response data were analyzed using PROAST v67.7 with treatment duration as a covariate. Data were obtained from Bhalli et al. (2013).

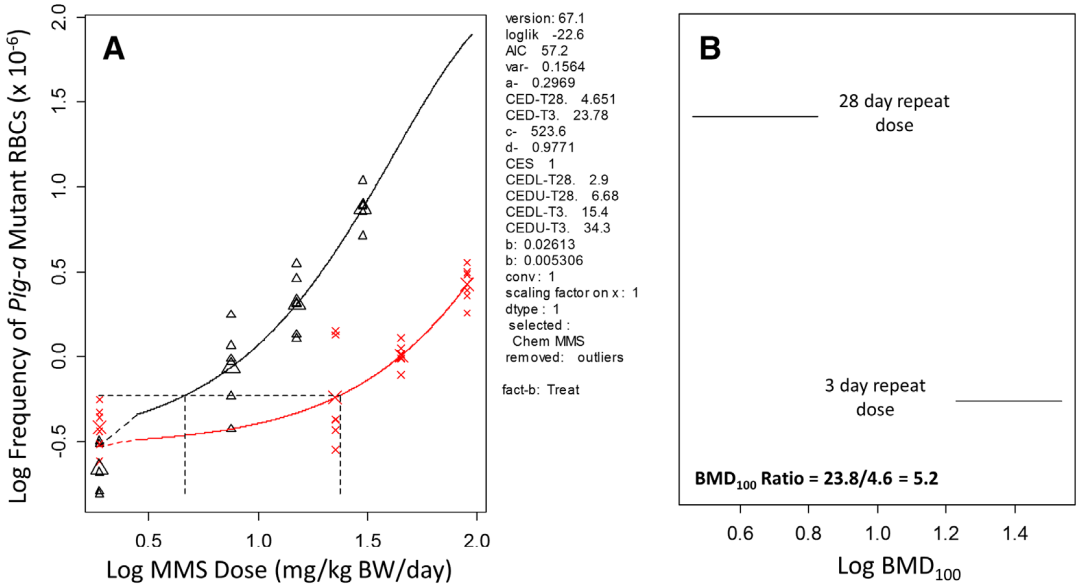

Fig. 4. The effect of treatment duration on the frequency of *Pig-a* mutant cells (i.e., CD59$^-$ cells). (A) Dose response relationship showing the frequency of *Pig-a* mutant rat RBCs (red blood cells) vs. oral dose of MMS. Red plotting symbols and fitted function show the results for the 3-day repeat-dose treatment, black shows the results for the 28-day repeat-dose treatment; post-exposure tissue sampling times are 12 and 14 days, respectively. Larger plotting symbols show mean values for each dose; smaller symbols show values for individual animals. Vertical dotted lines extending to the horizontal axis indicate BMD_{100} values (i.e., doubling dose). (**B**) Comparison of BMD_{100} values for the two treatment durations investigated. The left and right extremes of the horizontal lines indicate the lower and upper limits of the BMD 90% confidence interval, respectively. Dose–response data were analyzed using PROAST v67.7 with treatment duration as a covariate. Dose–response data are courtesy of Steve Dertinger, Litron Laboratories (Rochester, NY) (Dertinger et al. 2012).

Environmental and Molecular Mutagenesis. DOI 10.1002/em

discussion of issues concerning genotoxic effect severity is beyond the scope of this work and the reader is referred to the work of Heflich et al. (2019).

SUMMARY, CONCLUSIONS, AND FUTURE PROSPECTS

The paradigm shift in genetic toxicology from qualitative screen-and-bin to quantitative DR analysis is well underway; routine use of quantitative methods for interpretation of genotoxicity DR data is increasingly accepted. Likewise, there is increasing recognition that the BMD approach for determination of genotoxicity PoD values is robust and effective; it is preferable to the NOAEL since it is not restricted to one of the study doses, and a confidence limit can be readily calculated. Furthermore, there is a gradual movement toward the use of genotoxicity PoD values for risk assessment and regulatory decision-making. As noted, Johnson et al. and Muller and Gocke have already used *in vivo* mutagenicity DR data to determine PoD values from which HBGVs (i.e., TDI and PDE) can be calculated. More recently, some researchers have proposed the use of a similar approach to calculate PDE values for alkylnitrosamines that have been detected in therapeutic products. However, in all of these cases, the authors used AF values originally established for use with other toxicological endpoints. Therefore, although there is increasing support for quantitative interpretation of genotoxicity DR data for risk assessment, there is no empirical foundation underlying the AFs that have been used thus far, and consequently, an acute need to empirically determine robust genotoxicity AFs for calculation of genotoxicity HBGVs. As described herein, comparative analysis of DR data showing differential responses in wild-type and DNA repair-deficient cells/animals can be used for empirical determination of an intraspecies AF for human sensitivity (Figs. 1 and 2, Table I). Similar comparative analyses can be used to determine the need for AFs that account for interspecies TD differences, as well as AFs to more rigorously account for repeat-dose treatment durations that are less than half-lifetime (Figs. 3 and 4). With respect to interspecies TD differences specifically, future work could scrutinize differential sensitivity of human and animal cells to noteworthy genotoxicants (i.e., interspecies BMD ratios). With respect to study duration AFs, it will be necessary to generate DR data for extended *in vivo* treatment durations (e.g., >28 days). Additionally, as discussed in Heflich et al. (2019), there is also a need to judiciously consider issues related to the severity of genotoxic effects.

As noted, the MOE approach can be used to provide an indication of the risk posed by (geno)toxic substances. The approach, which simply involves calculation and evaluation of the ratio between a BMDL and the estimated or actual human exposure, provides a perspective on the level of concern for regulatory decision-making and risk management. The approach, which has been endorsed by the EFSA for regulatory decisions related to toxic contaminants in foods, employs the MOE approach to stipulate a level of concern. More specifically, the EFSA notes that MOE values greater than 10,000 are low concern for public health and a low priority for risk management; those less than 10,000 are used to indicate a necessity for further scrutiny and/or appropriate risk management (EFSA 2005). Although the EFSA has proposed the use of the BMD approach to calculate MOE values for substances that are both genotoxic and carcinogenic (EFSA 2005; Hardy et al. 2017), there are no standards to demarcate public health concerns for substances that are known to be genotoxic but hitherto not tested for carcinogenicity (Benford et al. 2010b; Benford 2016). If the MOE approach is to be used to prioritize regulatory action for genotoxic substances, the genetic toxicology community will need to establish MOE levels indicative of public health concern and the concomitant need for risk management (Benford 2016); specified levels may differ according to substance type and intended use (e.g., therapeutic products, pesticides, food contaminants, etc.).

The single, default AF values discussed earlier, which could be employed to determine genotoxicity HBGVs, effectively represent an upper bound estimate of the differences between species (i.e., interspecies AF for TD differences), differences among humans within a population (i.e., intraspecies AF), and compound potency (i.e., BMD) differences between short-term treatments and chronic treatments extending, for example, over half the lifetime of the experimental animal (i.e., study duration AF). As an alternative to single AF values, which are applied in a deterministic manner, it is possible to use the underlying uncertainty distributions to more effectively characterize hazard; moreover, combine AF uncertainty distributions for probabilistic hazard characterization. By characterizing and employing the uncertainty associated with each aspect of the hazard characterization, such an approach avoids the drawbacks associated with the deterministic approach (e.g., the assumption of conservatism) (Slob and Pieters 1998; Vermeire et al. 1999). Although the approach is more complex in comparison with that based on single deterministic values that are thought to be conservative, it has been simplified via the delineation of methods for approximate probabilistic analyses (IPCS 2014). By assuming that all AF uncertainties can be reflected by "independent lognormal probability distributions," approximate probabilistic analyses can effectively determine the human dose at which the risk of adverse effect is deemed to be negligible. To facilitate this type of analysis, an IPCS project has developed an Excel spreadsheet tool called APROBA (i.e. AProximate PROBabalistic Analysis) that can be used to determine the human dose associated with a particular magnitude of effect at a particular population incidence, as well as the uncertainty in that dose

Environmental and Molecular Mutagenesis. DOI 10.1002/em

(i.e., uncertainty in the HBGV) (IPCS 2014). Since it is critically important to consider uncertainties in risk assessment (i.e., uncertainties associated with HBGVs) (USEPA 2000), it is incumbent upon the genetic toxicology community to consider how such approaches might be used for determination of robust genotoxicity HBGVs. In the interim, the non-probabilistic approach can be pursued whereby only the upper and lower confidence limits on AFs are employed to determine HBGV confidence limits. More specifically, the distribution of PoD ratios such as those described earlier (Table I) can be employed to determine AF confidence limits, and by extension HBGV confidence limits. Nevertheless, rigorous use of upper and lower confidence limits on genetic toxicity AFs will necessitate eventual use of an approach such as APROBA for determination of robust genotoxicity HBGVs. To define AFs, and perhaps more importantly, AF distributions and confidence limits, future endeavors must expand on the analyses of published DR data presented herein.

It should also be noted that once appropriate endpoint-specific CES values have been established, mutagenicity BMD values can be empirically compared with carcinogenicity BMD values, and the relationships used to estimate carcinogenic potency (CP). This type of analysis has been conducted for the *in vivo* micronucleus (MN) endpoint (Hernández et al. 2012; Soeteman-Hernández et al. 2016), and the relationship between MN-inducing potency and CP recently used to estimate the CP of a feed additive metabolite (RIVM 2014). More generally, such analyses permit the use of *in vivo* mutagenicity DR data to estimate CP when those data are unavailable. Although the analyses conducted by Hernández et al. (2012) and Soeteman-Hernández et al. (2016) have already proved useful for a regulatory evaluation, the analyses were restricted to CP estimations based on BMD values for *in vivo* chromosome damage (i.e., MN induction). Thus, there is an acute need to extend this type of analysis to include *in vivo* mutagenicity endpoints such as TGR mutagenicity in selected tissues and *Pig-a* mutagenicity in peripheral blood. This regulatory use of mutagenicity DR data is particularly pertinent in light of (1) the high cost of carcinogenicity assays, (2) the acknowledged regulatory utility of CP values, and (3) global initiatives to reduce the use of experimental animals in toxicological testing (Beekhuijzen 2017; MacArthur 2018; Craig et al. 2019). With respect to regulatory utility, it is important to emphasize that for some types of substances (e.g., pesticides) in some jurisdictions, CP data are required for legislated regulatory evaluations (USC 1996; DOJ 2013).

Finally, it is important to emphasize that the risk assessment discussions presented herein only pertain to analysis and interpretation of *in vivo* DR data, that is, risk assessments based on *in vivo* PoD values. The utility of *in vitro* PoD values for risk assessment is controversial; it is currently being evaluated by several regulatory organizations (e.g., USEPA, Health Canada, ECHA, etc.). More specifically, toxicity assessment programs that employ high-throughput *in vitro* assays to generate extensive amounts of DR data (e.g., ToxCast and Tox21) are actively developing *in vitro* to *in vivo* extrapolation methods that permit the use of *in vitro* PoD values for human health risk assessment (Judson et al. 2010; Wetmore et al. 2012; Tice et al. 2013; Wetmore et al. 2015; Kavlock et al. 2018). A well-developed approach involves determination of an administered equivalent dose, such as an oral equivalent dose (OED), which corresponds to the steady-state blood level defined by an *in vitro* PoD. The OED is then compared with near- or far-field exposure estimates to determine a value that is analogous to an MOE, that is, the BER or bioactivity-exposure ratio (Wetmore et al. 2012, 2015; Kavlock et al. 2018). These values are used for risk-based prioritization; more specifically, determining the necessity for costly *in vivo* follow-up. Such an approach has never been employed for risk assessment of genotoxic substances, and future work should endeavor to employ a BER-based strategy for interpretation of *in vitro* genotoxicity PoD values. Interestingly, Health Canada, in collaboration with the USEPA and the US National Toxicology Program (NTP), is currently employing a suite of *in vitro* genotoxicity assays to generate DR data that will be used to evaluate the utility of *in vitro* genotoxicity DR data for risk assessment and regulatory decision-making (i.e., the GeneTox21 initiative). This type of work is congruent with global initiatives to improve the regulatory utility of high-throughput *in vitro* assays, thus reducing reliance on *in vivo* methods for hazard and risk assessment (Krewski et al. 2010; Thomas et al., 2013b; Kavlock et al. 2012).

ACKNOWLEDGMENTS

The authors gratefully acknowledge helpful comments and input from Christine Lemieux, Marc Beal, Julie Cox, and two anonymous reviewers. The work was supported by funding from the Government of Canada's Chemicals Management Plan.

AUTHOR CONTRIBUTIONS

The first draft of the manuscript was prepared by P.A.W. G.E.J. and A.S.L. provided substantial input for revision and preparation of the final version submitted for publication. All authors approved the final version prior to submission.

REFERENCES

Allan JM, Engelward BP, Dreslin AJ, Wyatt MD, Tomasz M, Samson LD. 1998. Mammalian 3-methyladenine DNA glycosylase

protects against the toxicity and clastogenicity of certain chemotherapeutic DNA cross-linking agents. Cancer Res 58:3965–3973.

Allemang A, Mahony C, Lester C, Pfuhler S. 2018. Relative potency of fifteen pyrrolizidine alkaloids to induce DNA damage as measured by micronucleus induction in HepaRG human liver cells. Food Chem Toxicol 121:72–81.

Altman DG, Bland JM. 1995. Absence of evidence is not evidence of absence. Br Med J 311:485.

Arimoto-Kobayashi S, Kaji K, Sweetman GM, Hayatsu H. 1997. Mutation and formation of methyl- and hydroxylguanine adducts in DNA caused by N-nitrosodimethylamine and N-nitrosodiethylamine with UVA irradiation. Carcinogenesis 18:2429–2433.

Barlow S, Renwick AG, Kleiner J, Bridges JW, Busk L, Dybing E, Edler L, Eisenbrand G, Fink-Gremmels J, Knaap A, et al. 2006. Risk assessment of substances that are both genotoxic and carcinogenic report of an International Conference organized by EFSA and WHO with support of ILSI Europe. Food Chem Toxicol 44:1636–1650.

Becker K, Thomas AD, Kaina B. 2014. Does increase in DNA repair allow "tolerance-to-insult" in chemical carcinogenesis? Skin tumor experiments with MGMT-overexpressing mice. Environ Mol Mutagen 55:145–150.

Beekhuijzen M. 2017. The era of 3Rs implementation in developmental and reproductive toxicity (DART) testing: Current overview and future perspectives. Reprod Toxicol 72:86–96.

Benford DJ. 2016. The use of dose-response data in a margin of exposure approach to carcinogenic risk assessment for genotoxic chemicals in food. Mutagenesis 31:329–331.

Benford D, Leblanc JC, Setzer RW. 2010a. Application of the margin of exposure (MoE) approach to substances in food that are genotoxic and carcinogenic: Example: Aflatoxin B1 (AFB1). Food Chem Toxicol 48(Suppl 1):S34–S41.

Benford D, Bolger PM, Carthew P, Coulet M, DiNovi M, Leblanc JC, Renwick AG, Setzer W, Schlatter J, Smith B, et al. 2010b. Application of the margin of exposure (MOE) approach to substances in food that are genotoxic and carcinogenic. Food Chem Toxicol 48 (Suppl 1):S2–S24.

Bhalli JA, Ding W, Shaddock JG, Pearce MG, Dobrovolsky VN, Heflich RH. 2013. Evaluating the weak in vivo micronucleus response of a genotoxic carcinogen, aristolochic acids. Mutat Res 753:82–92.

Bolger PM, Leblanc JC, Setzer RW. 2010. Application of the margin of exposure (MoE) approach to substances in food that are genotoxic and carcinogenic. Example: Acrylamide (CAS No. 79-06-1). Food Chem Toxicol 48(Suppl 1):S25–S33.

Bonassi S, El-Zein R, Bolognesi C, Fenech M. 2011. Micronuclei frequency in peripheral blood lymphocytes and cancer risk: Evidence from human studies. Mutagenesis 26:93–100.

Calabrese EJ, Gilbert CE. 1993. Lack of total independence of uncertainty factors (UFs): Implications for the size of the total uncertainty factor. Regul Toxicol Pharmacol 17:44–51.

Carthew P, DiNovi M, Setzer RW. 2010. Application of the Margin of Exposure (MOE) approach to substances in food that are genotoxic and carcinogenic. Example: PhIP (2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine) CAS No: 105650-23-5. Food Chem Toxicol 48(Suppl 1):S98–S105.

Cleaver JE, Feeney L, Tang JY, Tuttle P. 2007. Xeroderma pigmentosum group C in an isolated region of Guatemala. J Invest Dermatol 127:493–496.

COM (UK Committee on Mutagenicity). 2018. Statement on the Quantitative Approaches to the Assessment of Genotoxicity Data. Report 2018/S1:1-24. Didcot, Oxfordshire, United Kingdom: Public Health England, Chilton.

Coulet M, Leblanc JC, Setzer RW. 2010. Application of the margin of exposure (MoE) approach to substances in food that are genotoxic and carcinogenic: Example 12: Sudan I (CAS No. 842-07-9). Food Chem Toxicol 48 Suppl 1:S106–S111.

Craig E, Lowe K, Akerman G, Dawson J, May B, Reaves E, Lowit A. 2019. Reducing the need for animal testing while increasing efficiency in a pesticide regulatory setting: Lessons from the EPA Office of Pesticide Programs' Hazard and Science Policy Council. Regul Toxicol Pharmacol 108:104481.

Crump KS. 1984. An improved procedure for low-dose carcinogenic risk assessment from animal data. J Environ Pathol Toxicol Oncol 5: 339–348.

Daniels DS, Woo TT, Luu KX, Noll DM, Clarke ND, Pegg AE, Tainer JA. 2004. DNA binding and nucleotide flipping by the human DNA repair protein AGT. Nat Struct Mol Biol 11:714–720.

Dankovic DA, Naumann BD, Maier A, Dourson ML, Levy LS. 2015. The scientific basis of uncertainty factors used in setting occupational exposure limits. J Occup Environ Hyg 12(Suppl 1):S55–S68.

Dertinger SD, Phonethepswath S, Weller P, Avlasevich S, Torous DK, Mereness JA, Bryce SM, Bemis JC, Bell S, Portugal S, et al. 2011. Interlaboratory Pig-a gene mutation assay trial: Studies of 1,3-propane sultone with immunomagnetic enrichment of mutant erythrocytes. Environ Mol Mutagen 52:748–755.

Dertinger SD, Phonethepswath S, Avlasevich SL, Torous DK, Mereness J, Bryce SM, Bemis JC, Bell S, Weller P, Macgregor JT. 2012. Efficient monitoring of in vivo pig-a gene mutation and chromosomal damage: Summary of 7 published studies and results from 11 new reference compounds. Toxicol Sci 130:328–348.

Dertinger SD, Phonethepswath S, Avlasevich SL, Torous DK, Mereness J, Cottom J, Bemis JC, Macgregor JT. 2014. Pig-a gene mutation and micronucleated reticulocyte induction in rats exposed to tumorigenic doses of the leukemogenic agents chlorambucil, thiotepa, melphalan, and 1,3-propane sultone. Environ Mol Mutagen 55:299–308.

DNHW (Department of National Health and Welfare). 1993. The assessment of mutagenicity. Health protection branch mutagenicity guidelines. Environ Mol Mutagen 21:15–37.

Dobrovolsky VN, Elespuru RK, Bigger CA, Robison TW, Heflich RH. 2011. Monitoring humans for somatic mutation in the endogenous PIG-a gene using red blood cells. Environ Mol Mutagen 52:784–794.

DOJ (Department of Justice, Government of Canada). 2013. Pest Control Products Act (S.C. 2002, c. 28). Ottawa, Canada: Department of Justice.

ECHA (European Chemicals Agency). 2012. Guidance on Information Requirements and Chemical Safety Assessment. Chapter R.8: Charaterization of Dose[concentration]-response for Human Health. Report ECHA-2010-G-19-EN. Helsinki, Finland: ECHA.

ECHA (European Chemicals Agency). 2014. Guidance on Information Requirements and Chemical Safety Assessment. Chapter R.7a: Endpoint Specific Guidance, version 3.0. Helsinki, Finland: ECHA.

EFSA (European Food Safety Authority). 2005. Opinion of the Scientific Committee on a request from EFSA related to a harmonised approach for risk assessment of substances which are both genotoxic and carcinogenic. EFSA J 282:1–31.

EFSA (European Food Safety Authority). 2009. Use of benchmark dose approach in risk assessment: Guidance of the scientific committee. EFSA J 1150:1–72.

Engelward BP, Dreslin A, Christensen J, Huszar D, Kurahara C, Samson L. 1996. Repair-deficient 3-methyladenine DNA glycosylase homozygous mutant mouse cells have increased sensitivity to alkylation-induced chromosome damage and cell killing. EMBO J 15:945–952.

Engelward BP, Allan JM, Dreslin AJ, Kelly JD, Wu MM, Gold B, Samson LD. 1998. A chemical and genetic approach together define the biological consequences of 3-methyladenine lesions in the mammalian genome. J Biol Chem 273:5412–5418.

Environmental Mutagen Society Committee 17. 1975. Environmental mutagenic hazards. Science 187:503–514.

Erickson RP. 2010. Somatic gene mutation and human disease other than cancer: An update. Mutat Res 705:96–106.

Glassner BJ, Weeda G, Allan JM, Broekhof JL, Carls NH, Donker I, Engelward BP, Hampson, RJ, Hersmus, R, Hickman, MJ, Roth, RB, Warren HB, Wu, MM, Hoeijmakers, JH, Samson, LD. 1999. DNA repair methyltransferase (Mgmt) knockout mice are sensitive to the lethal effects of chemotherapeutic alkylating agents. Mutagenesis 14:339–347.

Godschalk R, Yauk C, van Benthem J, Douglas G, Marchetti F. 2020. In utero exposure to genotoxins leading to genetic mosaicism: An overlooked window of susceptibility in genetic toxicology testing? Environ Mol Mutagen 61:55–65.

Gollapudi BB, Johnson GE, Hernández LG, Pottenger LH, Dearfield KL, Jeffrey AM, Julien E, Kim JH, Lovell DP, Macgregor JT, et al. 2013. Quantitative approaches for assessing dose-response relationships in genetic toxicology studies. Environ Mol Mutagen 54:8–18.

Gollapudi BB, Lynch AM, Heflich RH, Dertinger SD, Dobrvolsky VN, Froetschl R, Horibata K, Kenyon MO, Kimoto T, Lovell DP, et al. 2015. The in vivo Pig-a assay: A report of the International Workshop on Genotoxicity Testing (IWGT) Workgroup. Mutat Res 783:23–35.

Goullet de Rugy T, Bashkurov M, Datti A, Betous R, Guitton-Sert L, Cazaux C, Durocher D, Hoffmann JS. 2016. Excess Poltheta functions in response to replicative stress in homologous recombination-proficient cancer cells. Biol Open 5:1485–1492.

Guerard M, Baum M, Bitsch A, Eisenbrand G, Elhajouji A, Epe B, Habermeyer M, Kaina B, Martus HJ, Pfuhler S, et al. 2015. Assessment of mechanisms driving non-linear dose-response relationships in genotoxicity testing. Mutat Res 763:181–201.

Guo X, Heflich RH, Dial SL, De M, Richter PA, Mei N. 2018. Quantitative differentiation of whole smoke solution-induced mutagenicity in the mouse lymphoma assay. Environ Mol Mutagen 59:103–113.

Haber LT, Maier A, Zhao Q, Dollarhide JS, Savage RE, Dourson ML. 2001. Applications of mechanistic data in risk assessment: The past, present, and future. Toxicol Sci 61:32–39.

Haber LT, Dourson ML, Allen BC, Hertzberg RC, Parker A, Vincent MJ, Maier A, Boobis AR. 2018. Benchmark dose (BMD) modeling: Current practice, issues, and challenges. Crit Rev Toxicol 48: 387–415.

Hardy A, Benford D, Halldorsson T, Jeger M, Knutsen K, More S, Mortensen A, Naegeli H, Noteburn H, Ockleford C, et al. 2017. Update: Use of the benchmark dose approach in risk assessment. EFSA J 15:4658.

Heflich RH, Johnson GE, Zeller A, Marchetti F, Douglas GR, Witt KL, Gollapudi BB, White PA. 2020. Mutation as a toxicological endpoint for regulatory decision-making. Environ Mol Mutagen 61:34–51.

Hernández LG, Van Benthem J, Slob W. 2012. Estimating the Carcinogenic Potency of Chemicals from the in vivo Micronucleus Test. RIVM Report 340700007/2012. Bilthoven, The Netherlands: RIVM (National Institute for Public Health and the Environment).

van der Hoeven JC, Lagerweij WJ, Bruggeman IM, Voragen FG, Koeman JH. 1983. Mutagenicity of extracts of some vegetables commonly consumed in The Netherlands. J Agric Food Chem 31: 1020–1026.

Horibata K, Ukai A, Ishikawa S, Sugano A, Honma M. 2016. Monitoring genotoxicity in patients receiving chemotherapy for cancer: Application of the PIG-A assay. Mutat Res 808:20–26.

IARC (International Agency for Research on Cancer). 2010a. IARC Monographs Working Group on the Evaluation of Carcinogenic Risks to Humans. Volume 95. Household Use of Solid Fuels and High-temperature Frying. IARC: Lyon, France.

IARC (International Agency for Research on Cancer). 2010b. IARC Monographs Working Group on the Evaluation of Carcinogenic Risks to Humans. Vol. 92. Some Non-heterocyclic Polycyclic Aromatic Hydrocarbons and Some Related Exposures. Lyon, France: IARC.

IARC (International Agency for Research on Cancer). 2014. IARC Monographs Working Group on the Evaluation of Carcinogenic Risks to Humans. Volume 105. Diesel and Gasoline Engine Exhausts and Some Nitroarenes. IARC: Lyon, France.

IARC (International Agency for Research on Cancer). 2016. IARC Monographs Working Group on the Evaluation of Carcinogenic Risks to Humans. Volume 109. Outdoor Air Pollution. Lyon, France: IARC.

ICH (International Council on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use). 2016. Impurities: Guidelines for Residual Solvents Q3C(R6). Geneva, Switzerland: ICH.

IPCS (International Programme on Chemical Safety). 1994. Assessing Human Health Risks of Chemicals: Derivation of Guidance Values for Health-Based Exposure Limits. Environmental Health Criteria 170. Geneva, Switzerland: World Health Organization.

IPCS (International Programme on Chemical Safety). 2014. Guidance Document on Evaluating and Expressing Uncertainly in Hazard Characterization. Harmonization Project Document 11. Geneva, Switzerland: World Health Organization.

IPCS (International Programme on Chemical Safety). 1999. Principles for the Assessment of risks to human health from exposure to chemicals. Environmental Health Criteria 210. Geneva, Switzerland: World Health Organization.

Jenkins GJ, Doak SH, Johnson GE, Quick E, Waters EM, Parry JM. 2005. Do dose response thresholds exist for genotoxic alkylating agents? Mutagenesis 20:389–398.

Johnson GE, Doak SH, Griffiths SM, Quick EL, Skibinski DO, Zair ZM, Jenkins GJ. 2009. Non-linear dose-response of DNA-reactive genotoxins: Recommendations for data analysis. Mutat Res 678: 95–100.

Johnson GE, Mikhail JM, Chapman K, Doak S, Seager AL, Shah K, Thomas A, Zair Z, Scott AD, Doherty A, et al. 2012. How do thresholds for mutagenicity and clastogenicity arise for DNA damaging agents? Genes Environ 34:179–185.

Johnson GE, Soeteman-Hernández L, Gollapudi BB, Bodger OG, Dearfield KL, Heflich RH, Hixon JG, Lovell DP, MacGregor JT, Pottenger LH, et al. 2014. Derivation of point of departure (PoD) estimates in genetic toxicology studies and their potential applications in risk assessment. Environ Mol Mutagen 55:609–623.

Judson RS, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, Mortensen HM, Reif DM, Rotroff DM, Imran S, Richard AM, et al. 2010. In vitro screening of environmental chemicals for targeted testing prioritization: The ToxCast Project. Environ Health Perspect 118:485–492.

Kavlock R, Chandler K, Houck K, Hunter S, Judson R, Kleinstreuer N, Knudsen T, Martin M, Padilla S, Reif D, et al. 2012. Update on EPA's ToxCast program: Providing high throughput decision support tools for chemical risk management. Chem Res Toxicol 25: 1287–1302.

Kavlock RJ, Bahadori T, Barton-Maclaren TS, Gwinn MR, Rasenberg M, Thomas RS. 2018. Accelerating the pace of chemical risk assessment. Chem Res Toxicol 31:287–290.

Kirsch-Volders M, Aardema M, Elhajouji A. 2000. Concepts of threshold in mutagenesis and carcinogenesis. Mutat Res 464:3–11.

Krewski D, Acosta D Jr, Andersen M, Anderson H, Bailar JC 3rd, Boekelheide K, Brent R, Charnley G, Cheung VG, Green S Jr, et al. 2010. Toxicity testing in the 21st century: A vision and a strategy. J Toxicol Environ Health B Crit Rev 13:51–138.

Li CY, Sung FC. 1999. A review of the healthy worker effect in occupational epidemiology. Occup Med (Lond) 49:225–229.

Long AS, Lemieux CL, Gagne R, Lambert IB, White PA. 2017. Genetic toxicity of complex mixtures of polycyclic aromatic hydrocarbons: Evaluating dose-additivity in a transgenic mouse model. Environ Sci Technol 51:8138–8148.

Lundstedt S, White PA, Lemieux CL, Lynes KD, Lambert IB, Oberg L, Haglund P, Tysklind M. 2007. Sources, fate, and toxic hazards of oxygenated polycyclic aromatic hydrocarbons (PAHs) at PAH-contaminated sites. AMBIO 36:475–485.

Lutz WK. 2009. The Viracept (nelfinavir)—ethyl methanesulfonate case: A threshold risk assessment for human exposure to a genotoxic drug contamination? Toxicol Lett 190:239–242.

MacArthur CJ. 2018. The 3Rs in research: A contemporary approach to replacement, reduction and refinement. Br J Nutr 120:S1–S7.

MacGregor JT. 2019. A Natural Mistake. Why Natural, Organic, and Botanical Products are Not as Safe as You Think. James T. MacGregor, Arnold, MD.

MacGregor JT, Frötschl R, White PA, Crump KS, Eastmond DA, Fukushima S, Guérard M, Hayashi M, Soeteman-Hernández LG, Johnson GE, et al. 2015a. IWGT report on quantitative approaches to genotoxicity risk assessment II. Use of point-of-departure (PoD) metrics in defining acceptable exposure limits and assessing human risk. Mutat Res 783:66–78.

MacGregor JT, Frotschl R, White PA, Crump KS, Eastmond DA, Fukushima S, Guerard M, Hayashi M, Soeteman-Hernández LG, Kasamatsu T, et al. 2015b. IWGT report on quantitative approaches to genotoxicity risk assessment I. Methods and metrics for defining exposure-response relationships and points of departure (PoDs). Mutat Res 783:55–65.

Marchetti F, Massarotti A, Yauk CL, Pacchierotti F, Russo A. 2016. The adverse outcome pathway (AOP) for chemical binding to tubulin in oocytes leading to aneuploid offspring. Environ Mol Mutagen 57:87–113.

Meier MJ, O'Brien JM, Beal MA, Allan B, Yauk CL, Marchetti F. 2017. In utero exposure to benzo[a]pyrene increases mutation burden in the soma and sperm of adult mice. Environ Health Perspect 125: 82–88.

Moreno NC, Garcia CCM, Munford V, Rocha CRR, Pelegrini AL, Corradi C, Sarasin A, Menck CFM. 2019. The key role of UVA-light induced oxidative stress in human Xeroderma Pigmentosum variant cells. Free Radic Biol Med 131:432–442.

Muller L, Gocke E. 2009. Considerations regarding a permitted daily exposure calculation for ethyl methanesulfonate. Toxicol Lett 190: 330–332.

Muller L, Gocke E, Lave T, Pfister T. 2009. Ethyl methanesulfonate toxicity in Viracept - A comprehensive human risk assessment based on threshold data for genotoxicity. Toxicol Lett 190:317–329.

OECD (Organization for Economic Cooperation and Development). 2013. OECD Guideline for the Testing of Chemicals, Test No. 488: Trangenic Rodent Somatic and Germ Cell Mutation Assays. Paris, France: OECD.

Pottenger LH, Schisler MR, Zhang F, Bartels MJ, Fontaine DD, McFadden LG, Bhaskar B. 2009. Dose-response and operational thresholds/NOAELs for in vitro mutagenic effects from DNA-reactive mutagens, MMS and MNU. Mutat Res 678:138–147.

Quinet A, Martins DJ, Vessoni AT, Biard D, Sarasin A, Stary A, Menck CF. 2016. Translesion synthesis mechanisms depend on the nature of DNA damage in UV-irradiated human cells. Nucleic Acids Res 44:5717–5731.

Renwick A, Leblanc JC, Setzer RW. 2010. Application of the margin of exposure (MoE) approach to substances in food that are genotoxic and carcinogenic - example: 1-methylcyclopropene and its impurities (1-chloro-2-methylpropene and 3-chloro-2-methylpropene). Food Chem Toxicol 48(Suppl 1):S81–S88.

Rey L, Sidorova JM, Puget N, Boudsocq F, Biard DS, Monnat RJ Jr, Cazaux, C, Hoffmann, JS. 2009. Human DNA polymerase eta is required for common fragile site stability during unperturbed DNA replication. Mol Cell Biol 29:3344–3354.

Ritter L, Totman C, Krishnan K, Carrier R, Vezina A, Morisset V. 2007. Deriving uncertainty factors for threshold chemical contaminants in drinking water. J Toxicol Environ Health B Crit Rev 10: 527–557.

RIVM (National Institute for Public Health and the Environment). 2014. Assessment of toxicity of furazolidone and 3-amino-2-oxazolidinone (AOZ). Report V/090130. Bilthoven, The Netherlands: RIVM.

Salk JJ, Schmitt MW, Loeb LA. 2018. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. Nat Rev Genet 19:269–285.

Shemansky JM, McDaniel LP, Klimas C, Dertinger SD, Dobrovolsky VN, Kimoto T, Horibata K, Polli JE, Heflich RH. 2019. Pig-a mutation database. Environ Mol Mutagen 60:759–762.

Slob W. 2002. Dose-response modeling of continuous endpoints. Toxicol Sci 66:298–312.

Slob W. 2016. A general theory of effect size, and its consequences for defining the benchmark response (BMR) for continuous endpoints. Crit Rev Toxicol 47(4):342–351.

Slob W, Pieters MN. 1998. A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: General framework. Risk Anal 18:787–798.

Slob W, Setzer RW. 2014. Shape and steepness of toxicological dose–response relationships of continuous endpoints. Crit Rev Toxicol 44:270–297.

Smith B, Cadby P, Leblanc JC, Setzer RW. 2010. Application of the margin of exposure (MoE) approach to substances in food that are genotoxic and carcinogenic: Example: Methyleugenol, CASRN: 93-15-2. Food Chem Toxicol 48(Suppl 1):S89–S97.

Smith-Roe SL, Swartz CD, Shepard KG, Bryce SM, Dertinger SD, Waidyanatha S, Kissling GE, Auerbach SS, Witt KL. 2018. Black cohosh extracts and powders induce micronuclei, a biomarker of genetic damage in human cells. Environ Mol Mutagen 59: 416–426.

Soeteman-Hernández LG, Johnson GE, Slob W. 2016. Estimating the carcinogenic potency of chemicals from the in vivo micronucleus test. Mutagenesis 31:347–358.

Stenson PD, Mort M, Ball EV, Evans K, Hayden M, Heywood S, Hussain M, Phillips AD, Cooper DN. 2017. The human gene mutation database: Towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. Hum Genet 136:665–677.

Thomas AD, Jenkins GJ, Kaina B, Bodger OG, Tomaszowski KH, Lewis PD, Doak SH, Johnson GE. 2013a. Influence of DNA repair on nonlinear dose-responses for mutation. Toxicol Sci 132:87–95.

Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, Devito MJ, Cote I, Rowlands JC, Whelan MP, Hays SM, Andersen ME, et al. 2013b. Incorporating new technologies into toxicity testing and risk

Environmental and Molecular Mutagenesis. DOI 10.1002/em

assessment: Moving from 21st century vision to a data-driven framework. Toxicol Sci 136:4–18.

Tice RR, Austin CP, Kavlock RJ, Bucher JR. 2013. Improving the human hazard characterization of chemicals: A Tox21 update. Environ Health Perspect 121:756–765.

Tomasetti C. 2019. Mutated clones are the new normal. Science 364: 938–939.

Torous DK. 2012. In vivo flow cytometric Pig-a and micronucleus assays: Highly sensitive discrimination of the carcinogen/noncarcinogen pair benzo(a)pyrene and pyrene using acute and repeated-dose designs. Environ Mol Mutagen 53:420–428.

USC (United States Code). 1996. The Federal Insecticide, Fungicide and Rodenticide Act, 7 U.S.C. ch. 6 § 136 et seq.

USEPA (United States Environmental Protection Agency). 2000. Science Policy Handbook: Risk Characterization. Report EPA 100-B-00-002. Washington, DC: USEPA.

USEPA (United States Environmental Protection Agency). 2012. Benchmark Dose Technical Guidance. Report EPA/100/R-12/001. Washington, DC: USEPA.

USEPA (United States Environmental Protection Agency). 2014. Guidance for Applying Quantitative Data to Develop Data-Derived Extrapolation Factors for Interspecies and Intraspecies Extrapolation. Report EPA/100/R-14/002F. Washington, DC: USEPA.

USFDA (United States Food and Drug Administration). 2005. Guidance for Industry. Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers. Rockville, MD: USFDA.

Vermeire T, Stevenson H, Peiters MN, Rennen M, Slob W, Hakkert BC. 1999. Assessment factors for human health risk assessment: A discussion paper. Crit Rev Toxicol 29:439–490.

Wetmore BA, Wambaugh JF, Ferguson SS, Sochaski MA, Rotroff DM, Freeman K, Clewell HJ 3rd, Dix DJ, Andersen ME, Houck KA, et al. 2012. Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. Toxicol Sci 125:157–174.

Wetmore BA, Wambaugh JF, Allen B, Ferguson SS, Sochaski MA, Setzer RW, Houck KA, Strope CL, Cantwell K, Judson RS, et al. 2015. Incorporating high-throughput exposure predictions with dosimetry-adjusted in vitro bioactivity to inform chemical toxicity testing. Toxicol Sci 148:121–136.

White PA, Johnson GE. 2016. Genetic toxicology at the crossroads-from qualitative hazard evaluation to quantitative risk assessment. Mutagenesis 31:233–237.

White PA, Zeller A, Pfuhler S, Johnson GE. 2019. Re: Gi et al. 2018, In vivo positive mutagenicity of 1,4-dioxane and quantitative analysis of its mutagenicity and carcinogenicity in rats, Archives of Toxicology 92:3207–3221. Arch Toxicol 93:211–212.

WHO (World Health Organisation). 2005. Chemical-Specific Adjustment Factors (CSAFs) for Interspecies Differences and Human Variability in Dose/Concentration-Response Assessment: Guidance Document for Use of Data in Dose/Concentration Response Assessment. International Programme on Chemical Safety (IPCS), Harmonization Project Document No. 2. Geneva, Switzerland: WHO.

WHO (World Health Organization). 2001. Guidance Document for the Use of Data in the Development of Chemical-Specific Adjustment Factors (CSAFs) for Interspecies Differences and Human Variability in Dose/Concentration-Response Assessment. Report WHO/PCS/01.4. Geneva, Switzerland: WHO.

Wignall JA, Shapiro AJ, Wright FA, Woodruff TJ, Chiu WA, Guyton KZ, Rusyn I. 2014. Standardizing benchmark dose calculations to improve science-based decisions in human health assessments. Environ Health Perspect 122:499–505.

Williams G, Leblanc JC, Setzer RW. 2010. Application of the margin of exposure (MoE) approach to substances in food that are genotoxic and carcinogenic: Example: (CAS No. 96-23-1) 1,3-dichloro-2-propanol (DCP). Food Chem Toxicol 48(Suppl 1):S57–S62.

Wills JW, Johnson GE, Doak SH, Soeteman-Hernández LG, Slob W, White PA. 2016a. Empirical analysis of BMD metrics in genetic toxicology part I: in vitro analyses to provide robust potency rankings and support MOA determinations. Mutagenesis 31:255–263.

Wills JW, Long AS, Johnson GE, Bemis JC, Dertinger SD, Slob W, White PA. 2016b. Empirical analysis of BMD metrics in genetic toxicology part II: in vivo potency comparisons to promote reductions in the use of experimental animals for genetic toxicity assessment. Mutagenesis 31:265–275.

Wills JW, Johnson GE, Slob W, White PA. 2017. Comparing BMD-derived genotoxic potency estimations across variants of the transgenic rodent gene mutation assay. Environ Molec Mutagen 58: 632–643.

Xu Y, Cui B, Ran R, Liu Y, Chen H, Kai G, Shi J. 2014. Risk assessment, formation, and mitigation of dietary acrylamide: Current status and future prospects. Food Chem Toxicol 69:1–12.

Yauk CL, Bishop J, Dearfield KL, Douglas GR, Hales BF, Luijten M, O'Brien JM, Robaire B, Sram R, van Benthem J, et al. 2013. The development of adverse outcome pathways for mutagenic effects for the organization for economic co-operation and development. Environ Mol Mutagen 54:79–81.

Yauk CL, Lambert IB, Meek ME, Douglas GR, Marchetti F. 2015. Development of the adverse outcome pathway "alkylation of DNA in male premeiotic germ cells leading to heritable mutations" using the OECD's users' handbook supplement. Environ Mol Mutagen 56:724–750.

Yebra-Pimentel I, Fernandez-Gonzalez R, Martinez-Carballo E, Simal-Gandara J. 2015. A critical review about the health risk assessment of PAHs and their metabolites in foods. Crit Rev Food Sci Nutr 55:1383–1405.

Youssoufian H, Pyeritz RE. 2002. Mechanisms and consequences of somatic mosaicism in humans. Nat Rev Genet 3:748–758.

Zair ZM, Jenkins GJ, Doak SH, Singh R, Brown K, Johnson GE. 2011. N-methylpurine DNA glycosylase plays a pivotal role in the threshold response of ethyl methanesulfonate-induced chromosome damage. Toxicol Sci 119:346–358.

Zeller A, Duran-Pacheco G, Guerard M. 2017. An appraisal of critical effect sizes for the benchmark dose approach to assess dose-response relationships in genetic toxicology. Arch Toxicol 91: 3799–3807.

Zhang L, Dong X, Lee M, Maslov AY, Wang T, Vijg J. 2019. Single-cell whole-genome sequencing reveals the functional landscape of somatic mutations in B lymphocytes across the human lifespan. Proc Natl Acad Sci USA 116:9014–9019.

Accepted by—
R. Heflich